



The Role of Artificial Intelligence and Machine Learning in Accelerating the Discovery and Development of Nanomedicine

Vivek Agrahari¹ · Yahya E. Choonara^{2,3} · Mitra Mosharraf^{4,5} · Sravan Kumar Patel^{6,7} · Fan Zhang^{8,9,10}

Received: 18 August 2024 / Accepted: 19 November 2024 / Published online: 2 December 2024
© The Author(s), under exclusive licence to Springer Science+Business Media, LLC, part of Springer Nature 2024

Abstract

The unique potential of nanomedicine to address challenging health issues is rapidly advancing the field, leading to the generation of more effective products. However, these complex systems often pose several challenges with respect to their design for specific functionality, scalable manufacturing, characterization, quality control, and clinical translation. In this regard, the application of artificial intelligence (AI) and machine learning (ML) approaches can enable faster and more accurate data assessment, identifying trends and predicting outcomes, leading to efficient nanomedicine product development. This perspective paper discusses the potential of AI and ML in nanomedicine product development with a focus on their applications in discovery, assessment, manufacturing, and clinical trials. The potential limitations of AI and ML approaches in nanomedicine product development are also covered.

Keywords artificial intelligence · machine learning · nanomedicine · nanoparticle · nanotechnology

Introduction

The field of nanomedicine is rapidly advancing and is poised for significant growth to support the development of health-care products. According to a market research report by Grand View Research, in the United States, the nanomedicine market was valued at USD 91.1 billion in 2023, encompassing nanoparticles (NPs), nanotubes, nanodevices and

nano shells [1]. NPs constitute the largest category (76%) of the market share [1]. These are generally classified as liposomes, lipid NPs (LNP), antibody–drug conjugates (ADCs), polymeric NPs, viral vectors, cell-derived NPs, inorganic NPs, nanocrystals, metal/metal oxide NPs, protein-based NPs, nanomicelles, and nanocrystals, etc.

Nanomedicines offer a versatile platform to deliver large array of therapeutics, including small molecule drugs, biologics, and vaccines. Particularly, due to their small size and favorable surface properties, nanomedicines can interact

Vivek Agrahari, Yahya E. Choonara, Mitra Mosharraf, Sravan Kumar Patel and Fan Zhang contributed equally.

✉ Mitra Mosharraf
mitra.mosharraf@htdcorp.com

¹ CONRAD, Eastern Virginia Medical School, Old Dominion University, Norfolk, VA 23507, USA

² Wits Advanced Drug Delivery Platform Research Unit, Department of Pharmacy and Pharmacology, School of Therapeutic Science, Faculty of Health Sciences, University of the Witwatersrand, Johannesburg, South Africa

³ Wits Infectious Diseases and Oncology Research Institute, Faculty of Health Science, University of the Witwatersrand, Johannesburg, South Africa

⁴ HTD Biosystems, 3197 Independence Drive, Livermore, CA 94551, USA

⁵ Engimata, 3197 Independence Drive, Livermore, CA 94551, USA

⁶ Department of Pharmaceutical Sciences, School of Pharmacy, University of Pittsburgh, Pittsburgh, PA 15213, USA

⁷ Magee-Womens Research Institute, Pittsburgh, PA 15213, USA

⁸ Department of Pharmaceutical Sciences, College of Pharmacy, University of Florida, 1350 Center Drive, Gainesville, FL 32610, USA

⁹ Department of Pharmacology & Therapeutics, College of Medicine, University of Florida, 1200 Newell Drive, Gainesville, FL 32610, USA

¹⁰ Department of Chemical Engineering, College of Engineering, University of Florida, 1006 Center Drive, Gainesville, FL 32611, USA

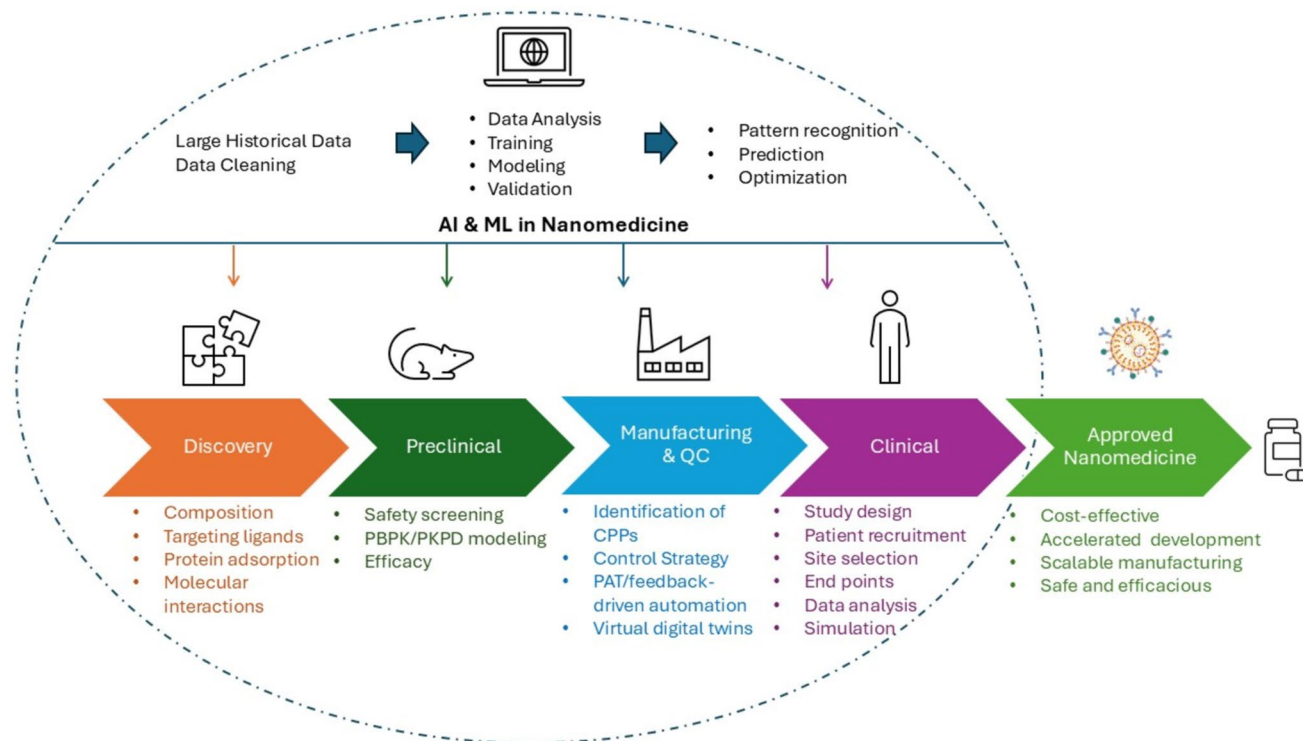


Fig. 1 A graphical illustration showing the integration of AI and ML in different stages of nanomedicine product development

with cells [2] and pass through complex physiological barriers such as the blood–brain barrier [3]. Therefore, these formulations have been widely investigated to address challenging health issues such as cancer, infectious diseases, and neurological disorders [3–5]. However, the major obstacles to the broader application of these products are their limited clinical translation, which stems from their poor *in vitro/in vivo* correlation, potential safety issues due to off-target effects, complexity and challenges in their scalable manufacturing, and product instability [6, 7].

The recent advances in Artificial Intelligence (AI) and Machine Learning (ML) and their potential contribution to healthcare product development offer a unique opportunity to integrate AI/ML at various stages of nanomedicine development. AI is a set of technologies that ultimately simulates human intelligence and as a result, it can perform complex tasks through learning, reasoning, and problem-solving [8]. ML is an area of computer science and AI, which utilizes historical data and past experiences to develop programs that can be used to model, predict, and control complex systems. Deep learning is a subset of ML that uses multilayered neural networks, called deep neural networks (DNN), to simulate the complex decision-making power of the human brain. Several reviews have been recently published that provided a concise incursion on the fundamentals of AI and ML algorithms [8–10]. For example, Kolluri and co-workers [9] focused on the

history of AI and ML and different tools that can be used for pharmaceutical applications. Usually, the AI/ML workflow starts with creating/accessing a large and diverse historical data library, the data is cleaned and organized, ensuring that the data labels and annotations are correct, the assumptions and correlational constraints are identified, and a model is selected. After data quality control and preparation, a portion of the available data (typically 70%) is used as the training set to train the model, while the remaining data is used for model verification and validation to prevent overfitting. In some cases, the model is validated by plotting the predicted values against experimental values to evaluate the model accuracy. The AI/ML workflow should include steps to ensure that the training data is of high quality to increase the success rate [10]. The AI/ML workflow in different stages of nanomedicine product development is illustrated in Fig. 1.

In this perspective paper, we have briefly discussed how AI and ML are leading the way in accelerating the discovery, manufacturing, pre-clinical, and clinical development of nanomedicines. We have also covered the potential limitations of AI/ML approaches in nanomedicine product development. The AI/ML tools help with screening and analyzing large data sets, modeling, prediction, process optimization, and decision making to ultimately reduce the time and cost of nanomedicine product development.

Application of AI in Nanomedicine Discovery and Design

AI has enabled efficient and systematic drug discovery for various therapeutic modalities ranging from small molecule drugs to biologics. However, it was not until the recent decade that AI has been harnessed for the design of nanomedicine formulations. Traditionally, the discovery of novel nanomedicine formulations relies on extensive trial and error. In the last decade, efforts have been made to apply AI for nanomedicine design, particularly in discovering new formulations that can achieve desirable properties such as the particle size distribution, drug loading, release kinetics, biodistribution, pharmacokinetics (PK), targeted delivery, safety, and efficacy.

To date, research in this direction has primarily focused on exploring large molecule drugs and excipients (e.g., lipids, polymers) to predict molecular interactions and identify novel nanomedicines with desirable properties using various deep learning algorithms. Specifically, Recurrent Neural Networks (RNNs), Convolutional Neural Networks (CNNs), Generative Adversarial Networks (GANs), and Transformers have been used to build ML platforms that can predict and achieve the desirable nanomedicine properties. For example, achieving high drug loading in nanoformulations has always been challenging, especially for hydrophobic drugs. Recent studies demonstrated ML-guided high-throughput experimentation to predict the self-assembly of hydrophobic drugs with either small molecular dyes or polymers [11–13]. These efforts led to the rapid and large-scale identification of nanomedicines with high drug-loading capacity. The use of AI and ML in long-acting nanomedicine formulation development is also emerging [14–19]. The accurate and rapid prediction of parameters such as long-term drug release, stability, and *in vivo* behavior, leveraged from AI-based imaging and ML algorithms have been reported, which can expedite the formulation development efforts [14]. Similarly, with Adeno-Associated Viruses (AAVs) for gene delivery, AI algorithms that can reduce the AAV discovery cycle by generating new libraries and selecting variants that increase the success rate of AAVs have been proposed [20].

One area that has reported significant progress on applying AI/ML based approaches is the formulation of messenger RNA (mRNA) LNPs. Historically, efforts to discover ionizable lipids for efficient mRNA delivery have relied on labor-intensive *in vitro* library screening. Li *et al.*, have recently developed an ML-based virtual screening model to identify novel ionizable lipids with high transfection efficiency [21]. Researchers have also used AI/ML-based approaches based on a synergistic

combination of deep learning and combinatorial chemistry to discover novel lipid molecules that would enable the design of LNPs to target specific cells or organs instead of liver and lymph nodes [22, 23]. It is worth noting that several centralized AI-based web platforms, such as FormulationAI [24] and AI-Guided Ionizable Lipid Engineering (AGILE) [23], have also been recently developed by integrating existing data on drugs and excipients. The AGILE is the first formulation-based platform that was built based on GNN algorithm to uniquely model the highly flexible yet straightforward molecular structures of ionizable lipids [23]. It has been used to predict novel ionizable lipids for the design and development of LNPs to deliver mRNA to muscle and immune cells [23]. For FormulationAI, multiple models based on more than 10 different algorithms were built for the selection of the best-fit model. Random Forest, Support Vector Machines (SVM), Light Gradient-Boosting Machine (LightGBM), and XGBoost were found to be the top-performing models [24]. FormulationAI enables the formulation design just by inputting basic information of drugs and excipients.

It is worth noting that due to the limited data availability, most of the work in this field, specifically the application of AI in formulation design, relies on in-house data generated from preformulation experiments or data derived from molecular computations (e.g., molecular descriptors, intermolecular distances and energies) [6–11]. For example, when training the centralized AGILE platform, researchers utilized internally generated data rather than widely accessible published data [11]. Of note, most of the published data associated with LNPs *in vitro* performance is based on heat maps or bar graphs, and often lacks a standard control, making it challenging for AI scientists to perform data processing and comparisons. This highlighted a critical need for scientists to consider the meta-analysis when releasing the study results. Equally important is the data quality, which often has a great impact on the quality of the model. A healthy unbiased model can be achieved through data audits and evaluations to ensure that the underlying training data is diverse, representative, and free from biases.

The application of AI in nanomedicine design can also be extended beyond the design of excipients such as novel lipids and toward prediction of drug-exipient interactions. AlfaFold [25] and AlfaFold2 [26] can efficiently predict the protein or peptide structures and their interactions, which can accelerate the discovery of protein or peptide-based targeting moieties or therapeutic payload. These payloads and targeting moieties can be then used as building blocks in the design of novel nanomedicine products. A recent study used AI/ML to screen peptide structures that can efficiently penetrate the cells and bind to the target [27]. Integration of these structures into nanomedicine design with AI-assisted

peptide engineering can generate formulations with high targeting efficiency.

Application of AI in Nanomedicine Preclinical Pharmacology

Understanding the *in vivo* pharmacology of nanomedicines is time- and labor-intensive and clinical translation often do not align with preclinical discovery, given the complicated interaction of nanomedicines with body components. Recently, progress in this direction has been made by integrating AI/ML with nanomedicines *in vivo* studies. Through analyzing existing data such as PK, biodistribution, efficacy, toxicity, and NP protein corona, researchers can identify the underlying relations that help predict the *in vivo* fate of nanomedicines. Physiologically Based Pharmacokinetic (PBPK) modeling of nanomedicines is at the forefront of this area [28]. PBPK modeling uses a combination of anatomical, physiological, and the analyte's physicochemical, transport, and elimination properties to predict the drug *in vivo* exposure. Given the need for a large number of experiments to obtain input data, AI-based approaches can generate virtual data from literature and using simulations, which can then be validated using experimental data. When combined with the quantitative structure–activity relationship (QSAR) model connecting the nanomedicine *in vivo* activity (e.g. tumor targeting) with its physicochemical properties (e.g. size, surface charge), AI-assisted PBPK models can guide the optimization of nanomedicines to achieve desirable tissue targeting properties and minimize toxicity [29, 30]. For example, Chou et. al, collated the physiochemical properties of NPs from more than 200 publications as the input features in an AI-assisted QSAR model (Random Forests and DNN) [29]. This model predicts critical tumor-related kinetics parameters (e.g. tumor deposition, release, clearance rate) that are essential to construct the AI-assisted PBPK model. The work provided an efficient screening tool to rapidly predict the delivery efficiency of NPs based on its molecular fingerprints such as size, charge, shape, etc., which can facilitate the clinical translation of nanomedicine [29]. Another recent study used ML-based models to identify critical design parameters for inorganic NPs to achieve efficacy against tumors in animal models [31].

ML models are also trained to analyze complex microscopy data and to differentiate between biological structures. A large amount of high-resolution imaging data on the *in vivo* biodistribution of NPs has helped researchers train ML-based models to advance the understanding of NP behavior *in vivo*, such as the mechanism of NP interaction with tumor vasculature [32]. Kingston and co-workers [17] used deep learning for automatic image segmentation of tumor tissue images to enable accurate staging. Specifically, manually

annotated images were used for the training and establishment of a segmentation model based on the U-net convolutional neural network. This model was then used for automatic machine segmentation of tumor tissue images with high accuracy and efficiency [32]. In another example, images of NP deposition in tumor metastasis were used to train SVM models to predict NPs distribution in metastatic tumors. The quadratic SVM model was found to have the best performance [17].

Generative AI has also gained significant momentum to develop novel customizable patient-centric nanomedicines [18, 19]. For example, Tang and co-workers developed a Generative Adversarial Network for Distribution Analysis (GANDA) to describe the intratumoral distribution of quantum dots post-injection [33]. The model was developed using tumor vessels and cell nuclei decomposed from whole-slide images of breast cancer sections to generate images of identical spatial resolution with minimal loss and reliability.

NPs biological behavior such as protein corona formation and opsonization, which influences their therapeutic efficacy, is significantly affected by their surface characteristics. Numerous factors (quantitative and qualitative) can influence protein corona formation, making it difficult to predict the protein corona using a general linear regression model or density functional theory. A recent study built a model by integrating ML and meta-analysis [34], providing an effective method to predict the functional compositions of protein corona and the cell recognition of NPs. Lazarovits et. al. [35] used experimental proteomics data of NPs to develop a supervised DNN model to assess liver/spleen accumulation from protein adsorption on NPs and predict their *in vivo* behavior. Researchers have also developed ML algorithms to design NPs with high drug loading and/or to deliver drugs to specific cells for broad therapeutic applications [11, 36].

Another application of AI/ML that can significantly de-risk nanomedicine translation is prediction and optimization of safety profile. For example, these tools are used for analyzing large amount of bioanalytical data obtained from characterization of biomarkers (e.g. anti-drug antibodies) from patient samples to predict immunogenicity or tolerance [37].

Application of AI in Nanomedicine Manufacturing, Characterization, and Quality Control

Application of AI-driven methods such as deep learning techniques have advanced nanomedicine manufacturing and quality control due to their ability to analyze large unstructured data and their superior predictability compared to traditional methods. Applications reported in the literature primarily focused on formulation development, identifying

critical quality attributes (CQAs) and critical process parameters (CPPs), process optimization, and quality control through process analytical technology (PAT). Rebollo *et al.* applied design of experiments (DoE) and artificial neural networks (ANNs) to identify critical formulation and process parameters to rationally design liposomes with an optimized particle size and polydispersity index (PDI) [38]. Most influential factors such as the flow rate were used as inputs for the ANN model to predict particle size and PDI of liposomes. The resilient backpropagation with weight backtracking ANN was developed. Notably, data from both screening and response surface methods were used to build the ANN model, demonstrating the ability of AI methods to handle data not generated from predetermined experimental designs. The ANN approach enabled the development of a unified model for PDI and size, which is not feasible with a DoE-based approach.

CQAs such as particle size, surface charge, and the composition of NPs can significantly influence their *in vivo* activity. Hence, reproducible production of NPs is critical, though challenging, as small changes in process parameters can have a significant impact on their CQAs. In this respect, integrating cutting-edge analytical technologies with AI-based data analysis and ML algorithms that provide real-time analysis can significantly improve the reproducibility of manufacturing processes. Real-time analysis allows for immediate identification of potential defects and deviations from desired CQAs, which enables better control and automation of the manufacturing process in a feedback-driven approach. This leads to consistent robust manufacturing, ultimately improving final product quality. AI can also advance the understanding and control of CPPs that affect the CQAs of nanomedicines to minimize the risk of product quality-related failure. Several AI and ML-based approaches including ANNs, deep learning, SVM, genetic algorithms, and multiple linear regression have been applied in nanomedicine manufacturing and characterization [39, 40]. For example, a genetic algorithm-derived image analysis has been used to analyze electron microscopy images of NPs and identify manufacturing process defects/impurities [41]. The method analyzed more than 150,000 NPs with a high precision rate of 99.75%. The potential of ML models in deriving gold NPs size/shape using DLS and UV–vis data is also another example of applications of AI/ML tools in NP's CQA testing [42].

AI and ML tools have also been increasingly applied within the area of continuous manufacturing. The PAT initiative in pharmaceutical continuous manufacturing aims at real-time measurements of CQAs and CPPs using in-line, at-line, and on-line sensors to enable correcting and preventing errors, adjusting process conditions to avoid a failed batch, ultimately leading to an agile manufacturing paradigm. PAT tools such as near-infrared and Raman spectroscopy enable

real-time monitoring of product CQAs and impurities, leading to improved control of the manufacturing process and product quality. The integration of PAT and continuous manufacturing leads to a large amount of multivariate data. In this regard, ML-based approaches such as ANNs, which are inherently flexible and can adjust model behavior for continuous performance improvement, can analyze big data in real-time allowing for adaptive control, autonomous decision-making, and the creation of digital twins and smart factories [43]. PAT tools (mostly based on light-scattering methods) for in-line measurement of NP characteristics have been reported. For example, in-line size (NanoFlowSizer) and residual ethanol measurements were applied in a continuous liposomal manufacturing process to control CQAs of the product through automated feedback and correction of process parameters such as mixing rate [24].

AI and ML-based algorithm models have been explored in several studies [39, 40, 44, 45] to optimize the continuous manufacturing process and CQAs of liposome formulations. For example, Sansare *et al.* [40] developed an ANN model to accurately predict the CQAs of liposomes (particle size and PDI). The model included critical material attributes (CMAs), CPPs, and molecular descriptors of lipids as inputs, and CQAs obtained from an at-line particle sizer, as outputs. Of the two ANN architectures used, the Multi-Input Single-Output (MISO) model constituting particle size in one neural network and PDI in another outperformed the Multi-Input Multi-Output (MIMO) model. In developing the model, training and testing errors were monitored. The optimized model that included molecular descriptors showed a lower mean relative error between predicted and actual values. The findings indicated that algorithmic models can effectively predict the CQAs of liposomes and other nanoformulations. Such AI/ML-based studies are highly recommended not only to generate development platforms but also to reduce the time and resources required for nanoformulations optimization, ultimately expediting their clinical development process.

To improve prediction and allow automation of PAT approaches, identification of product defects and out-of-specifications, and support real-time monitoring of product physicochemical properties, the pattern recognition and prediction power of AI can have significant applications in nanomedicine manufacturing. For example, the integration of advanced data analytic tools powered by AI can significantly improve the performance and reliability of PAT, as reported for tablets [43]. Another application of AI is to improve the model performance of process digital twins by providing virtual data points, for instance of manufacturing processes, which enables investigation of various process conditions without the need for physical experimentations [46, 47]. Davidopoulou *et al.* [47] have combined ANNs with a digital twin process for milling operation to predict

and optimize process outcomes of nanosuspension production. Specifically, ANN was incorporated within data sampling, model deployment, and curve fitting to augment the experimental data set and simulate the dynamic comminution profile inside the physical mill, which can estimate process outcomes such as particle size profile at any time point during milling. Overall, while still in the early stages, AI and ML-based approaches combined with PAT will be indispensable tools to monitor, control, and optimize a multitude of factors in real time to ensure the reproducible manufacturing of nanomedicines.

Application of AI in Nanomedicine Clinical Trials

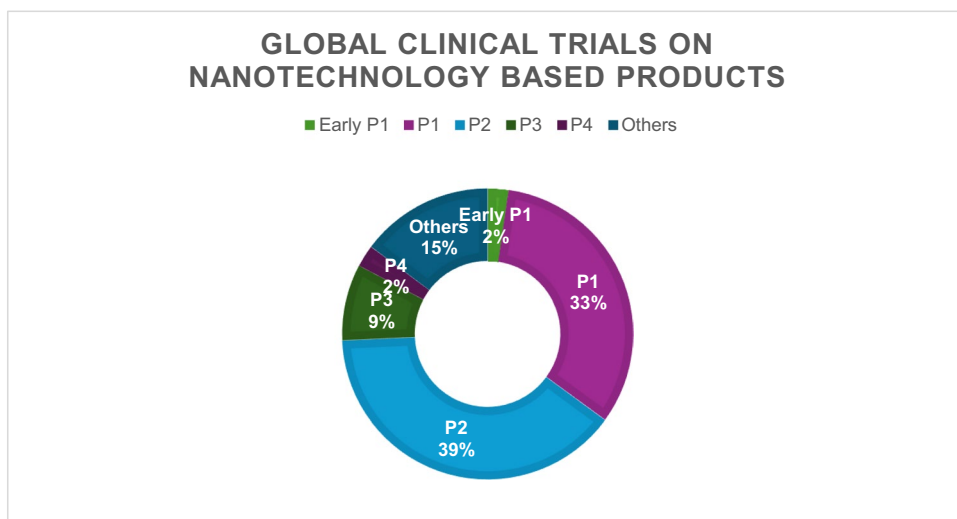
According to ClinicalTrials.gov, from January 2024 to June 2024, there were 660 nanotechnology-related global clinical trials. As shown in Fig. 2, 34% were in phase 1, 40% in phase 2, 9% in phase 3, and 2% in phase 4. Clinical trials are costly and time-consuming, with the overall success rate across all therapeutics being approximately 10.8% [48]. Advances in AI and ML have revolutionized clinical trials by accelerating processes (e.g. increasing patient recruitment rates), simulating trials, and predicting outcomes [9], thereby potentially increasing the clinical trial success rates and reducing the costs and time to market. AI has also enabled pharmaceutical companies to better understand drug safety and efficacy [49, 50]. By managing and analyzing large datasets, AI facilitates pattern recognition and optimization in clinical trials.

Key areas benefiting from AI integration in clinical trials include patient identification, recruitment, management, retention, dose selection, and trial design [51]. During the

COVID-19 pandemic, several pharmaceutical companies used AI to expedite nanotechnology-based COVID-19 vaccine development [52]. ML and historical data were employed for disease forecasting and participant recruitment, to predict outcomes and to optimize clinical trials [53]. Smart Data Query (SDQ), an ML tool for accelerating data review, was used to review clinical trial data for a NP-based mRNA vaccine in 22 h, a process that traditionally took a month [53]. Such approaches have also helped leveraging AI to develop a platform that predicts patient responses to both approved and novel therapies. This platform integrates over a trillion data points to optimize clinical trials, allowing pharmaceutical and biotech companies to rapidly conduct thousands of trial simulations. AI-based simulation models can be used to decide upon the percent of each patient group that should be included in a clinical trial. This significantly increases the probability of trial success and accelerates drug development while minimizing risks. In general, there are many advantages of using AI in clinical trials. AI can be used to analyze large sets of patient data or electronic health records to identify patients' characteristics that are important for the response or study design. It can identify and recruit patients who are most likely to benefit from a particular treatment and identify subjects that are at high risk of study non-compliance and increase compliance by remote monitoring in real-time. Simulating patient recruitment scenarios, clinical trial design, and outcome, are other areas where AI is used to help reduce the time and cost of recruiting patients for clinical trials [51].

AI can also be used to collect, manage, and analyze large clinical trial datasets to identify patterns and trends that can help us understand drug safety and efficacy in ways that would not be possible without ML tools. For example, Lee *et al.* reported creating an AI/ML system called SEETrials [50], allowing a large language model to extract safety and

Fig. 2 The breakdown of global clinical trials on nanotechnology-based products from January 2024 to June 2024. Data was extracted from clinicaltrials.gov. The graph was generated based on the percentage of nanotechnology-based products in clinical trials at each stage of development (early P1, P2, P3, and P4 during this period. Category “others” includes studies that are not in P1 to P4 clinical trials)



efficacy data from clinical trial abstracts, from key conferences and PubMed across different years (2012–2023) to predict safety and efficacy of clinical trials in oncology [54]. They used four modules: pre-processing, prompt modeling, knowledge ingestion and post-processing, and evaluated the system's performance both qualitatively and quantitatively. In addition, they evaluated the model across different cancer types (multiple myeloma (MM), breast, lung, lymphoma, and leukemia) and studied the efficacy and safety of CAR-T, bispecific antibodies, and ADCs, across a large scale of clinical trial studies. They demonstrated that SEETrials revealed highly accurate data extraction across different cancer types and therapeutic areas showing high precision (0.958), recall (sensitivity) (0.944), and F1 score (0.951) across 70 data elements present in the MM trial studies.

Another area that can leverage AI in clinical trials data collection is remote patient monitoring using digital health technologies (e.g., wearable devices and digital biomarkers) without requiring the patients to visit a health facility for sample collection. Automated data collection, surveillance, and verification can help reduce errors and improve quality. In addition, clinical trial protocols can be optimized by analyzing data from prior trials [51]. AI can also help with designing more effective and efficient trials. For example, identifying the optimal dose for a drug and selecting the most appropriate patients for a particular treatment [50].

The inclusion of AI in the clinical trials market is rapidly expanding, with a compound annual growth rate (CAGR) of 28.6%, projected to reach \$22.3 billion by 2033 [55]. However, when using ML tools to design and manage clinical trials, it is important to ensure that the source data used for training AI/ML models are diverse and free from bias [56].

Limitations of AI/ML in Nanomedicine Product Development

The use of AI and ML-based approaches in the development of nanomedicine products has led to a significant reduction in the time and cost required for their discovery, manufacturing, assessing PK, and clinical design and testing. Despite their several advantages, AI/ML-based models also have some limitations [57–59]. Benders and Cortes-Ciriano [10] have discussed the limitations of AI in drug discovery and highlighted the importance of data quality for predictions and success rates. When using AI/ML models, it is important to ensure that the underlying training data is large, diverse, of high quality, and free from bias. If the training data is small and not diverse enough, it can lead to limited predictive power and bias. There is a potential to miss a perfectly fit clinical candidate because the data may not fit the pattern that the system has been trained on.

To prevent biases, it is important that the algorithm is based on accurate assumptions and correlation constraints. In clinical testing, ML algorithms should regularly evaluate and address bias because these can lead to false predictions and inconsistent performance. There is a risk that patients from underrepresented groups such as women, minorities and from low- and middle-income countries may not be sufficiently represented in the clinical trial design. Sponsors must inform patients about the risks and benefits of data sharing through AI/ML models, ensuring patient consent and retention throughout trials, given the ethical and security issues. In addition, employing AI/ML approaches necessitates use of computers and data storage capabilities. Lack of access to such infrastructures and cost constraints could limit organizations from adapting the AI/ML tools in the product development process.

To reduce the risk of using inaccurate algorithms and to reduce discriminatory outcomes, it is also of utmost importance to have clear regulatory guidelines regarding ML-driven data monitoring. To address these concerns, the U.S. Food and Drug Administration (FDA) has recently published two discussion papers, one on using AI and ML in the development of drug and biological products [60], and one on AI in drug manufacturing [61]. FDA is planning to release guidance on using AI in drug development in 2024 [62].

Conclusions

AI is poised to play a pivotal role in surmounting some of the key challenges of nanomedicines to better predict and optimize their discovery process, development, scalable manufacturing, and *in vivo* pharmacology before clinical trials. Several ML-based tools can specifically enhance the PK/PD properties of NPs to facilitate their clinical translation by using the large amount of PK and biodistribution data available in literature. In addition, the reproducible manufacturing of nanomedicines is critical and hence AI-integrated PAT can improve process understanding that positively impacts the CQAs from a Quality by Design (QbD) perspective that is integral to regulatory success. The FDA is proactively developing its regulatory framework for AI to ensure that AI integration in product development and clinical trials is both effective and ethical. Overall, integrating AI and ML tools with nanomedicine development will make it more efficient, cost-effective, and translatable for broader clinical applications. However, when leveraging novel AI-enabled modalities, it is important to ensure the model is validated for accurate prediction of nanomedicine characteristics. Incorporating a process for data quality control is also essential. Such mitigation strategies can be applied to overcome biases caused by ML enabled predictive models in the

pharmaceutical industry across discovery, manufacturing, and clinical development. Overall, to bridge the gaps in integrating AI/ML-based approaches in nanomedicine research, effective collaboration among product development scientists, ML engineers, data scientists, robotics engineers, regulatory agencies, AI researchers, and clinicians is needed to ensure reproducible and constructive use of this fascinating technology in product development approaches.

Acknowledgements All authors are part of the American Association of Pharmaceutical Scientists (AAPS) Nanotechnology Community leadership team. The authors would like to acknowledge the support of other members in the AAPS leadership team. F.Z.'s research effort reported in this publication was supported by the National Center for Advancing Translational Sciences of the National Institutes of Health (NIH) under Award Number 5KL2TR001429-09. The content is solely the responsibility of the authors and does not necessarily represent the official views of the NIH.

Data Availability The data that supports the findings of Fig. 2 (breakdown of global clinical trials on nanotechnology-based products from January 2024 to June 2024) was extracted from clinicaltrials.gov. The data included the number of trials in different phases of development (early P1, P2, P3, and P4) during this period, and “others” (i.e. studies that are not in P1 to P4 clinical trials). These data were used to calculate the percentage of nanotechnology-based products in clinical trials at each stage of development. The data that supports the findings of this study are available from the corresponding author upon reasonable request.

Declarations

Conflict of Interest All authors certify that they have no conflict of interest with the subject matter or materials discussed in this manuscript.

Disclaimer The views expressed in this perspective review article are solely those of the authors and do not reflect the views of the author's employer, company, or institution.

References

1. U.S. Nanomedicine Market Size, Share & Trends Analysis Report By Application (Drug Delivery), By Indication (Clinical Oncology, Infectious Diseases), By Molecule Type, By Region, And Segment Forecasts, 2024 - 2030. Grand View Research.
2. Thapa RK, Kim JO. Nanomedicine-based commercial formulations: current developments and future prospects. *J Pharm Investig.* 2023;53(1):19–33.
3. Hersh AM, Alomari S, Tyler BM. Crossing the Blood-Brain Barrier: Advances in Nanoparticle Technology for Drug Delivery in Neuro-Oncology. *Int J Mol Sci.* 2022;23(8):4153.
4. Mir Najib Ullah SN, *et al.* Nanomedicine in the Management of Alzheimer's Disease: State-of-the-Art. *Biomedicines.* 2023;11(6):1752.
5. Kirtane AR, *et al.* Nanotechnology approaches for global infectious diseases. *Nat Nanotechnol.* 2021;16(4):369–84.
6. Agrahari V, Agrahari V. Facilitating the translation of nanomedicines to a clinical product: challenges and opportunities. *Drug Discov Today.* 2018;23(5):974–91.
7. Dordevic S, *et al.* Current hurdles to the translation of nanomedicines from bench to the clinic. *Drug Deliv Transl Res.* 2022;12(3):500–25.
8. Suriyaamporn P, *et al.* The artificial intelligence-powered new era in pharmaceutical research and development: A review. *AAPS PharmSciTech.* 2024;25(6):188.
9. Kolluri S, *et al.* Machine learning and artificial intelligence in pharmaceutical research and development: A review. *AAPS J.* 2022;24(1):19.
10. Bender A, Cortés-Ciriano I. Artificial intelligence in drug discovery: what is realistic, what are illusions? Part 1: Ways to make an impact, and why we are not there yet. *Drug Disc Today.* 2021;26(2):511–24.
11. Reker D, *et al.* Computationally guided high-throughput design of self-assembling drug nanoparticles. *Nat Nanotechnol.* 2021;16(6):725–33.
12. Shamay Y, *et al.* Quantitative self-assembly prediction yields targeted nanomedicines. *Nat Mater.* 2018;17(4):361–8.
13. Alves VM, *et al.* Cheminformatics-driven discovery of polymeric micelle formulations for poorly soluble drugs. *Sci Adv.* 2019;5(6):eaav9784.
14. Bannigan P, *et al.* Machine learning models to accelerate the design of polymeric long-acting injectables. *Nat Commun.* 2023;14(1):35.
15. Clark AG, *et al.* Investigating structural attributes of drug encapsulated microspheres with quantitative X-ray imaging. *J Control Release.* 2023;358:626–35.
16. Clark AG, *et al.* Assessing microstructural critical quality attributes in PLGA microspheres by FIB-SEM analytics. *J Control Release.* 2022;349:580–91.
17. Kingston BR, *et al.* Assessing micrometastases as a target for nanoparticles using 3D microscopy and machine learning. *Proc Natl Acad Sci U S A.* 2019;116(30):14937–46.
18. Ghebrehiwet I, *et al.* Revolutionizing personalized medicine with generative AI: a systematic review. *Artif Intell Rev.* 2024;57(5):1–41.
19. Bannigan P *et al.* The Dawn of a New Pharmaceutical Epoch: Can AI and Robotics Reshape Drug Formulation? *Advanced Healthcare Materials*, 2024: p. 2401312. <https://doi.org/10.1002/adhm.202401312>.
20. Ail D, *et al.* Adeno-Associated Virus (AAV) - Based Gene Therapies for Retinal Diseases: Where are We? *Appl Clin Genet.* 2023;16:111–30.
21. Li B, *et al.* Accelerating ionizable lipid discovery for mRNA delivery using machine learning and combinatorial chemistry. *Nat Mater.* 2024;23(7):1002–8.
22. Sahay G. The Quest for Next Generation Lipid Nanoparticles for Cell-Specific Delivery of mRNA. in *AAPS National Biotechnology Conference.* 2024.
23. Xu Y, *et al.* AGILE platform: a deep learning powered approach to accelerate LNP development for mRNA delivery. *Nat Commun.* 2024;15(1):6305.
24. Dong J *et al.* FormulationAI: a novel web-based platform for drug formulation design driven by artificial intelligence. *Brief Bioinform.* 2023;25(1). <https://doi.org/10.1093/bib/bbad419>.
25. Jumper J, *et al.* Highly accurate protein structure prediction with AlphaFold. *Nature.* 2021;596(7873):583–9.
26. Yang Z, *et al.* AlphaFold2 and its applications in the fields of biology and medicine. *Signal Transduct Target Ther.* 2023;8(1):115.
27. Hsueh HT, *et al.* Machine learning-driven multifunctional peptide engineering for sustained ocular drug delivery. *Nat Commun.* 2023;14(1):2509.
28. Yuan D, *et al.* Physiologically based pharmacokinetic modeling of nanoparticles. *J Pharm Sci.* 2019;108(1):58–72.

29. Chou WC, *et al.* An artificial intelligence-assisted physiologically-based pharmacokinetic model to predict nanoparticle delivery to tumors in mice. *J Control Release.* 2023;361:53–63.
30. Singh AV, *et al.* Artificial intelligence and machine learning disciplines with the potential to improve the nanotoxicology and nanomedicine fields: a comprehensive review. *Arch Toxicol.* 2023;97(4):963–79.
31. Mendes BB, *et al.* A large-scale machine learning analysis of inorganic nanoparticles in preclinical cancer research. *Nat Nanotechnol.* 2024;19(6):867–78.
32. Zhu M, *et al.* Machine-learning-assisted single-vessel analysis of nanoparticle permeability in tumour vasculatures. *Nat Nanotechnol.* 2023;18(6):657–66.
33. Tang Y, *et al.* GANDA: A deep generative adversarial network conditionally generates intratumoral nanoparticles distribution pixels-to-pixels. *J Control Release.* 2021;336:336–43.
34. Ban Z, *et al.* Machine learning predicts the functional composition of the protein corona and the cellular recognition of nanoparticles. *Proc Natl Acad Sci U S A.* 2020;117(19):10492–9.
35. Lazarovits J, *et al.* Supervised Learning and Mass Spectrometry Predicts the in Vivo Fate of Nanomaterials. *ACS Nano.* 2019;13(7):8023–34.
36. Boehnke N, *et al.* Massively parallel pooled screening reveals genomic determinants of nanoparticle delivery. *Science.* 2022;377(6604):eabm5551.
37. Duhaze J, *et al.* A machine learning approach for high-dimensional time-to-event prediction with application to immunogenicity of biotherapies in the ABIRISK cohort. *Front Immunol.* 2020;11:608.
38. Rebollo R, *et al.* Microfluidic manufacturing of liposomes: Development and optimization by design of experiment and machine learning. *ACS Appl Mater Interfaces.* 2022;14(35):39736–45.
39. Di Francesco V, *et al.* Machine learning instructed microfluidic synthesis of curcumin-loaded liposomes. *Biomed Microdevices.* 2023;25(3):29.
40. Sansare S, *et al.* Artificial neural networks in tandem with molecular descriptors as predictive tools for continuous liposome manufacturing. *Int J Pharm.* 2021;603:120713.
41. Lee B, *et al.* Statistical characterization of the morphologies of nanoparticles through machine learning based electron microscopy image analysis. *ACS Nano.* 2020;14(12):17125–33.
42. Glaubitz C, *et al.* Leveraging machine learning for size and shape analysis of nanoparticles: A shortcut to electron microscopy. *J Phys Chem C Nanomater Interfaces.* 2024;128(1):421–7.
43. Nagy B, *et al.* Application of artificial neural networks in the process analytical technology of pharmaceutical manufacturing—a review. *AAPS J.* 2022;24(4):74.
44. Hoseini B, *et al.* Application of ensemble machine learning approach to assess the factors affecting size and polydispersity index of liposomal nanoparticles. *Sci Rep.* 2023;13(1):18012.
45. Han R, *et al.* Predicting liposome formulations by the integrated machine learning and molecular modeling approaches. *Asian J Pharm Sci.* 2023;18(3):100811.
46. Malheiro V, *et al.* Exploiting Pharma 4.0 Technologies in the Non-Biological Complex Drugs Manufacturing: Innovations and Implications. *Pharmaceutics.* 2023;15(11):2545.
47. Davidopoulou C, Ouranidis A. Pharma 4.0-Artificially Intelligent Digital Twins for Solidified Nanosuspensions. *Pharmaceutics.* 2022;14(10):2113.
48. Global Trends in R&D 2024: Activity, productivity, and enablers. 2024: IQVIA.
49. Harrer S, *et al.* Artificial intelligence for clinical trial design. *Trends Pharmacol Sci.* 2019;40(8):577–91.
50. Hutson M. How AI is being used to accelerate clinical trials. *Nature.* 2024;627(8003):S2–5.
51. Weissler EH, *et al.* Correction to: The role of machine learning in clinical research: transforming the future of evidence generation. *Trials.* 2021;22(1):593.
52. Sharma A, *et al.* artificial intelligence-based data-driven strategy to accelerate research, development, and clinical trials of COVID vaccine. *Biomed Res Int.* 2022;2022:7205241.
53. Pfizer.com, mRNA and Artificial Intelligence for Advanced Vaccine Innovation. Available from: https://www.pfizer.com/news/articles/how_a_novel_incubation_sandbox_helped_speed_up_data_analysis_in_pfizer_s_covid_19_vaccine_trial. Accessed 11/21/2024.
54. Lee K *et al.* SEETrials: Leveraging Large Language Models for Safety and Efficacy Extraction in Oncology Clinical Trials. *medRxiv*, 2024: p. 2024.01. 18.24301502.
55. Market.us extracted from Clinicaltrials.gov. Available from: <https://market.us/report/ai-in-clinical-trial-market/>. Accessed 11/21/2024.
56. 2 Key Areas To Leverage AI/ML For More Successful Clinical Trials. [cited 2024 August 04]; Available from: <https://www.clinicalleader.com/doc/key-areas-to-leverage-ai-ml-for-more-successful-clinical-trials-0001>. Accessed 11/21/2024.
57. Villa Nova M, *et al.* Nanomedicine ex machina: between model-informed development and artificial intelligence. *Front Digit Health.* 2022;4:799341.
58. Das KP. Nanoparticles and convergence of artificial intelligence for targeted drug delivery for cancer therapy: Current progress and challenges. *Front Med Technol.* 2022;4:1067144.
59. Naik GG, Jagtap VA. Two Heads Are Better than One: Unraveling the Potential Impact of Artificial Intelligence in Nanotechnology. *Nano TransMed.* 2024;3:100041.
60. Administration, U.S.F.D. Using Artificial Intelligence & Machine Learning in the Development of Drug & Biological Products - Discussion Paper and Request for Feedback. 2023 [cited 2024 November 02]; Available from: <https://www.fda.gov/media/167973/download>. Accessed 11/21/2024.
61. Center for Drug Evaluation and Research, U.S.F.D.A. Artificial Intelligence in Drug Manufacturing. 2023 [cited 2024 November 01]; Available from: <https://www.fda.gov/media/165743/download>. Accessed 11/21/2024.
62. FDA plans to release AI drug development guidance this year. 2024 [cited 2024 August 11]; Available from: <https://www.raps.org/news-and-articles/news-articles/2024/5/fda-plans-to-release-ai-drug-development-guidance>. Accessed 11/21/2024.

Publisher's Note Springer Nature remains neutral with regard to jurisdictional claims in published maps and institutional affiliations.

Springer Nature or its licensor (e.g. a society or other partner) holds exclusive rights to this article under a publishing agreement with the author(s) or other rightsholder(s); author self-archiving of the accepted manuscript version of this article is solely governed by the terms of such publishing agreement and applicable law.